1. The oldest business in the world

Image: St. Peter Stiftskeller, founded 803. Credit: Pakeha (https://commons.wikimedia.org/wiki/File:Eingang_zum_St._Peter_Stiftskeller.jpg).

An important part of business is planning for the future and ensuring that the company survives changing market conditions. Some businesses do this really well and last for hundreds of years.

BusinessFinancing.co.uk researched (https://businessfinancing.co.uk/the-oldest-company-in-almost-every-country) the oldest company that is still in business in (almost) every country and compiled the results into a dataset. In this project, you'll explore that dataset to see what they found.

The database contains three tables.

`categories`

column	type	meaning
category_code	varchar	Code for the category of the business.

column	type	meaning
category	varchar	Description of the business category

countries

column	type	meaning
country_code	varchar	ISO 3166-1 3-letter country code.
country	varchar	Name of the country.
continent	varchar	Name of the continent that the country exists in.

businesses

column	type	meaning
business	varchar	Name of the business.
year_founded	int	Year the business was founded.
category_code	varchar	Code for the category of the business.
country_code	char	ISO 3166-1 3-letter country code.

Let's begin by looking at the range of the founding years throughout the world.

```
In [175]:  %%sql
           postgresql:///oldestbusinesses
           select  min(year_founded), max(year_founded)
           from businesses
```

1 rows affected.

Out[175]:

min	max
578	1999

2. How many businesses were founded before 1000?

Wow! That's a lot of variation between countries. In one country, the oldest business was only founded in 1999. By contrast, the oldest business in the world was founded back in 578. That's pretty incredible that a business has survived for more than a millennium.

I wonder how many other businesses there are like that.

```
In [177]:  %%sql
           select count(*)
           from businesses
           where year_founded < 1000
```

```
 * postgresql:///oldestbusinesses
1 rows affected.
```

Out[177]:

count
6

3. Which businesses were founded before 1000?

Having a count is all very well, but I'd like more detail. Which businesses have been around for more than a millennium?

In [179]:
```
%%sql
select * from businesses
where year_founded < 1000
order by year_founded
```

 * postgresql:///oldestbusinesses
6 rows affected.

Out[179]:

business	year_founded	category_code	country_code
Kongō Gumi	578	CAT6	JPN
St. Peter Stifts Kulinarium	803	CAT4	AUT
Staffelter Hof Winery	862	CAT9	DEU
Monnaie de Paris	864	CAT12	FRA
The Royal Mint	886	CAT12	GBR
Sean's Bar	900	CAT4	IRL

4. Exploring the categories

Now we know that the oldest, continuously operating company in the world is called Kongō Gumi. But was does that company do? The category codes in the `businesses` table aren't very helpful: the descriptions of the categories are stored in the `categories` table.

This is a common problem: for data storage, it's better to keep different types of data in different tables, but for analysis, you want all the data in one place. To solve this, you'll have to join the two tables together.

In [181]:
```
%%sql
select business, year_founded, businesses.country_code, category
from businesses
inner join categories on businesses.category_code = categories.category_code
where year_founded < 1000
order by year_founded
```

 * postgresql:///oldestbusinesses
6 rows affected.

Out[181]:

business	year_founded	country_code	category
Kongō Gumi	578	JPN	Construction
St. Peter Stifts Kulinarium	803	AUT	Cafés, Restaurants & Bars
Staffelter Hof Winery	862	DEU	Distillers, Vintners, & Breweries
Monnaie de Paris	864	FRA	Manufacturing & Production
The Royal Mint	886	GBR	Manufacturing & Production
Sean's Bar	900	IRL	Cafés, Restaurants & Bars

5. Counting the categories

With that extra detail about the oldest businesses, we can see that Kongō Gumi is a construction company. In that list of six businesses, we also see a café, a winery, and a bar. The two companies recorded as "Manufacturing and Production" are both mints. That is, they produce currency.

I'm curious as to what other industries constitute the oldest companies around the world, and which industries are most common.

In [183]:
```sql
%%sql
select category, count(*) as n
from businesses
inner join categories on businesses.category_code = categories.category_code
group by category
order by n desc
limit 10
```

 * postgresql:///oldestbusinesses
10 rows affected.

Out[183]:

category	n
Banking & Finance	37
Distillers, Vintners, & Breweries	22
Aviation & Transport	19
Postal Service	16
Manufacturing & Production	15
Media	7
Agriculture	6
Cafés, Restaurants & Bars	6
Food & Beverages	6
Tourism & Hotels	4

6. Oldest business by continent

It looks like "Banking & Finance" is the most popular category. Maybe that's where you should aim if you want to start a thousand-year business.

One thing we haven't looked at yet is where in the world these really old businesses are. To answer these questions, we'll need to join the `businesses` table to the `countries` table. Let's start by asking how old the oldest business is on each continent.

In [185]:
```sql
%%sql
select min(year_founded) as oldest, continent
from businesses
inner join countries on businesses.country_code = countries.country_code
group by continent
order by oldest
```

 * postgresql:///oldestbusinesses
6 rows affected.

Out[185]:

oldest	continent
578	Asia
803	Europe
1534	North America
1565	South America
1772	Africa
1809	Oceania

7. Joining everything for further analysis

Interesting. There's a jump in time from the older businesses in Asia and Europe to the 16th Century oldest businesses in North and South America, then to the 18th and 19th Century oldest businesses in Africa and Oceania.

As mentioned earlier, when analyzing data it's often really helpful to have all the tables you want access to joined together into a single set of results that can be analyzed further. Here, that means we need to join all three tables.

In [187]:
```sql
%%sql
select business, year_founded, category, country, continent
from businesses
inner join categories using(category_code)
inner join countries using(country_code)
```

 * postgresql:///oldestbusinesses
163 rows affected.

Out[187]:

business	year_founded	category	country	continent
Spinzar Cotton Company	1930	Agriculture	Afghanistan	Asia
ALBtelecom	1912	Telecommunications	Albania	Europe
Andbank	1930	Banking & Finance	Andorra	Europe
Liwa Chemicals	1939	Manufacturing & Production	United Arab Emirates	Asia
Bank of the Province of Buenos Aires	1822	Banking & Finance	Argentina	South America
Yerevan Ararat Brandy-Wine-Vodka Factory	1877	Distillers, Vintners, & Breweries	Armenia	Asia
Australia Post	1809	Postal Service	Australia	Oceania
St. Peter Stifts	803	Cafés, Restaurants &	Austria	Europe

8. Counting categories by continent

Having `businesses` joined to `categories` and `countries` together means we can ask questions about both these things together. For example, which are the most common categories for the oldest businesses on each continent?

In [189]:
```sql
%%sql
select continent, category, count(*) as n
from categories
inner join businesses using(category_code)
inner join countries on businesses.country_code = countries.country_code
group by continent, category
```

 * postgresql:///oldestbusinesses
56 rows affected.

Out[189]:

continent	category	n
North America	Banking & Finance	4
Oceania	Postal Service	1
South America	Food & Beverages	2
Europe	Tourism & Hotels	2
Asia	Media	1
Europe	Medical	1
Asia	Defense	1
Africa	Manufacturing & Production	1
Europe	Postal Service	4
North America	Aviation & Transport	2
Asia	Distillers, Vintners, & Breweries	2
South America	Banking & Finance	3
North America	Food & Beverages	1
Europe	Manufacturing & Production	8
Africa	Postal Service	9
Asia	Telecommunications	1
Africa	Food & Beverages	1
Europe	Consumer Goods	3
Europe	Mining	1
Oceania	Banking & Finance	2
Asia	Agriculture	1
North America	Manufacturing & Production	1
Africa	Mining	1
Africa	Aviation & Transport	10
Asia	Construction	2
Asia	Energy	3
Asia	Retail	3
South America	Manufacturing & Production	2
Asia	Cafés, Restaurants & Bars	3
Europe	Banking & Finance	5
Africa	Banking & Finance	17
North America	Tourism & Hotels	1
Europe	Cafés, Restaurants & Bars	2
Asia	Banking & Finance	6
South America	Defense	1
Africa	Energy	1
Asia	Conglomerate	3
North America	Media	1
Europe	Agriculture	1

Asia	Mining	1
Asia	Aviation & Transport	7
Asia	Food & Beverages	2
North America	Distillers, Vintners, & Breweries	5
Africa	Agriculture	3
Africa	Distillers, Vintners, & Breweries	3
North America	Agriculture	1
Asia	Postal Service	2
Europe	Defense	1
Asia	Manufacturing & Production	3
Europe	Telecommunications	1
Europe	Distillers, Vintners, & Breweries	12
South America	Cafés, Restaurants & Bars	1
North America	Retail	1
Africa	Media	4
Europe	Media	1
Asia	Tourism & Hotels	1

9. Filtering counts by continent and category

Combining continent and business category led to a lot of results. It's difficult to see what is important. To trim this down to a manageable size, let's restrict the results to only continent/category pairs with a high count.

In [191]:
```sql
%%sql
select continent, category, count(*) as n
from categories
inner join businesses using(category_code)
inner join countries on businesses.country_code = countries.country_code
group by continent, category
having count(*) > 5
order by n desc
```

 * postgresql:///oldestbusinesses
7 rows affected.

Out[191]:

continent	category	n
Africa	Banking & Finance	17
Europe	Distillers, Vintners, & Breweries	12
Africa	Aviation & Transport	10
Africa	Postal Service	9
Europe	Manufacturing & Production	8
Asia	Aviation & Transport	7
Asia	Banking & Finance	6